Study
Note
97-01

Virtual Environment Interface Requirements for Combat Leader Training and Mission Rehearsal

Paul J. Sticha
Roy C. Campbell
Human Resources Research Organization

Steven R. Schwalm
Hughes Training, Inc.

United States Army Research Institute for the Behavioral and Social Sciences

August 1997

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract for the Department of the Army

Human Resources Research Organization

Technical review by

Bob G. Witmer

NOTICES

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) 1997, August	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) August 1995 - May 1996
4. TITLE AND SUBTITLE Virtual Environment Interface Requirements for Combat Leader Training and Mission Rehearsal		5a. CONTRACT OR GRANT NUMBER DASW01-94-0011 (Delivery Order 05) 5b. PROGRAM ELEMENT NUMBER 0605803A
6. AUTHOR(S) Paul J. Sticha and Roy C. Campbell (HumRRO); and Steven R. Schwalm (Hughes Training, Inc.)		5c. PROJECT NUMBER 2O665803D730 5d. TASK NUMBER 2111 5e. WORK UNIT NUMBER C01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization (HumRRO) 66 Canal Center Plaza, Suite 400 Alexandria, Virginia 22314		8. PERFORMING ORGANIZATION REPORT NUMBER Final Report (FR-WATSD-96-15)
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-IF 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Study Note 97-01

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: Bruce W. Knerr

14. ABSTRACT *(Maximum 200 words)*:

The Army has made a substantial commitment to the use of networked computer simulations for training, concept development, and test and evaluation. The current networked training system -- Simulation Networking (SIMNET) -- and the next generation system -- the Close Combat Tactical Trainer (CCTT) -- provide effective forms of training for soldiers fighting from vehicles, but these systems are unable to do the same for individual dismounted soldiers. Virtual Environment (VE) technology has the potential to provide Individual Combat Simulations (ICS) for the electronic battlefield.

This report reviews the current state-of-the-art and projected future capabilities of the VE technologies associated with speech recognition, gesture recognition, and computer-generated forces. The review provides a roadmap that outlines the potential applications of these VE technologies for training, mission rehearsal, and performance measurement for combat team leaders; enumerates the technological capabilities need to implement these applications; specifies realistic near-term goals for prototype development and testing; and identifies knowledge gaps and the research needed to fill them.

15. SUBJECT TERMS Virtual Environment Voice Recognition Gesture Recognition Computer-Generated Forces

SECURITY CLASSIFICATION OF					
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	19. LIMITATION OF ABSTRACT Unlimited	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)

Study Note 97-01

Virtual Environment Interface Requirements for Combat Leader Training and Mission Rehearsal

Paul J. Sticha
Roy C. Campbell
Human Resources Research Organization

Steven R. Schwalm
Hughes Training, Inc.

Simulator Systems Research Unit
Stephen L. Goldberg, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

August 1997

Army Project Number
20665803D730

Personnel and Training Analysis Activities

Preceding Page Blank

FOREWORD

The U.S. Army has made a substantial commitment to the use of networked computer simulations for training, concept development, and test and evaluation. The current networked training system—Simulation Networking (SIMNET)—and the next generation system—the Close Combat Tactical Trainer (CCTT)—provide effective forms of training for soldiers fighting from vehicles, but these systems are unable to do the same for individual dismounted soldiers. Virtual Environment (VE) technology has the potential to provide Individual Combat Simulations (ICS) for the electronic battlefield.

One of the most promising potential applications of VE is training and mission rehearsal for the small combat unit leader (platoon, squad, or fire team). Because these leaders interact directly with their subordinates, it places especially severe demands on those technologies that permit direct interaction with and control of computer-controlled subordinates. When fully developed and integrated, these technologies will permit training the small unit leader in combat decision-making, communication, and leadership skills in a realistic combined arms environment without the necessity of equipping an entire squad of soldiers with expensive VE interfaces.

This report reviews the current state-of-the-art and projected future capabilities in the component VE technologies associated with speech recognition, gesture recognition, and computer-generated forces. The review provides a roadmap that outlines the potential applications of these VE technologies for training, mission rehearsal, and performance measurement for combat team leaders; enumerates the technological capabilities need to implement these applications; specifies realistic near-term goals for prototype development and testing; and identifies knowledge gaps and the research needed to fill them.

ZITA M. SIMUTIS Technical Director	EDGAR M. JOHNSON Director

VIRTUAL ENVIRONMENT INTERFACE REQUIREMENTS FOR COMBAT LEADER TRAINING AND MISSION REHEARSAL

CONTENTS

Page

LIST OF TABLES

VIRTUAL ENVIRONMENT INTERFACE REQUIREMENTS FOR COMBAT LEADER TRAINING AND MISSION REHEARSAL

Introduction

The U.S. Army has made a substantial commitment to the use of networked computer simulations for combined arms training, concept development, and test and evaluation. The current networked training system, Simulation Networking (SIMNET), and its successor system, the Close Combat Tactical Trainer (CCTT), can provide effective training for units employing a variety of combat vehicles. The Multiservice Distributed Training Testbed (MDT2) incorporates air assets into these simulations, allowing units from different Services and at different sites to coordinate in performing Close Air Support (CAS) missions.

Despite the advances in providing distributed training and mission rehearsal capabilities to soldiers in combat vehicles, current systems are not able to provide these capabilities to dismounted soldiers. The large visual field of view, the importance of relatively small terrain features, and the use of verbal commands and hand-and-arm signals for communications present significant challenges to simulation technology. Recent developments in Virtual Environment (VE) technology have potential to meet some of the training and mission rehearsal needs of dismounted soldiers.

The Defense Modeling and Simulation Office (DMSO) and the U. S. Army Research Institute for the Behavioral and Social Sciences (ARI) have jointly funded research to develop a prototype system that applies VE technologies to provide a more realistic interface for the dismounted squad or fire team leader. This research has focused on the following technologies: (a) speech recognition, (b) gesture recognition, and (c) computer-generated forces (CGF). This research has combined off-the-shelf hardware and software with additional, specially designed software. Substantial progress has been made on these research efforts. Consequently, there is now a need to identify training strategies that could take advantage of VE capabilities for training and mission rehearsal for combat team leaders. Furthermore, the specific tasks for which these strategies would be employed and the relevant training settings need to be specified. Finally, research must be specified to resolve the uncertainties regarding interface requirements and to provide the guidance required for implementation of devices using VE technologies for training and mission rehearsal.

The goals of this project, therefore, are to provide a roadmap that: (a) outlines the potential applications of VE technology incorporating speech recognition, gesture recognition, and CGF to training, mission rehearsal, and performance measurement for combat team leaders; (b) numerates the technological capabilities needed to implement these applications; (c) specifies realistic near-term goals for prototype development and testing; and (d) identifies knowledge gaps and the research needed to fill them.

To meet these goals, we conducted several activities. First, we identified the combat leader tasks in which training, mission planning and rehearsal, and performance assessment could be enhanced by the use of speech recognition, gesture recognition, and CGF. We then surveyed the

current and projected capabilities of relevant VE technologies to meet the needs presented by the leader tasks. A comparison of requirements and capabilities then identified areas that require technological development and areas that require additional behavioral research. The behavioral research is required to evaluate the performance of alternative technologies, to determine the effect that technological limitations have on system effectiveness, and to address instructional strategies and overall effectiveness issues.

Applications of VE Technology for Combat Leaders

The focus of this report is on the small unit leader. This individual could be a squad or fire team leader, or the leader of a Special Forces Detachment. This level was chosen for several reasons. First, the small unit leader must react to factors such as the terrain, enemy disposition, and the condition of his own troops to lead his unit to accomplish its mission. Such skills are well-suited to training using interactive simulation. However, small, dismounted units such as squads are currently not represented directly in DIS, and the potential application of DIS is largely unmet. It may never be economically sound to equip an entire squad for training in VE. However, it may be sound to equip the leader for training in VE, and use CGF to simulate his squad members. Consequently, the small unit leader is a good candidate for development of simulation-based training using VE technology.

Although we consider training to be the primary application of VE technology for the small unit leader, we envision applications of this technology in mission planning and rehearsal, performance assessment, and concept development. Training applications could vary in complexity and skill level required. At one end of the spectrum, VE training could give the squad leader a chance to practice simple skills, such as giving hand-and-arm signals and issuing verbal commands, in a wide variety of situations. This practice would give the squad leader a breadth of experience that is not possible in other training methods. However, we expect a greater training value would arise from the use of VE technology to train more advanced skills. These skills could be of roughly the complexity of ARTEP Battle Drills, or they could be more integrated missions. The VE system could serve for both training delivery and performance assessment.

A second application of VE is for mission planning and rehearsal. Benefits for mission planning would be greatest if the system had a terrain data base matching the actual terrain on which the mission was to be performed. Initially, the squad leader could use VE technology to perform mission and terrain analysis. During this phase, he could try out and evaluate several concepts of operations. He could then use the technology to brief the plan to members of his unit or to higher echelons. Finally, the leader and/or members of his unit could rehearse critical parts of the plan in a realistic setting.

A third application of VE technology is as an aid to the concept development process. VE technology could be used to evaluate the effects of new equipment, doctrine, or organizational structure on a unit's capability to accomplish its mission. Equipment capabilities that do not currently exist could be simulated in VE to determine their potential value for increasing combat effectiveness. The information obtained in this manner could be used to modify the planned equipment, organization, or doctrine. Individual VE systems would be used in this area in

conjunction with larger-scale interactive simulations and constructive simulations. For example, an individual VE simulation might be used to obtain parameters for a human performance model that would then be incorporated into a constructive simulation.

Organization of This Report

The next section of this report describes some representative combat leader activities that are candidates for training using VE technology. Individual activities are combined into nine training scenarios that provide practice on a variety of skills in a realistic setting. We briefly describe each scenario and discuss some of the considerations for performing them in a simulated setting.

The discussion of leader activities is followed by a description of the capabilities in the three VE technologies that are the subject of this report. We describe the capabilities of existing, off-the-shelf systems, and project future capabilities. Comparing these capabilities to the requirements indicates which scenarios can be implemented in the short term, and which will require further technological developments.

The following section describes the functional requirements of the combat leader activities in terms of the features of the relevant VE technologies. These requirements will be approximate for some activities and technologies, because further research is required to determine what level of technology is required to perform or train some activities in VE. We will point out uncertainties that are reasonable topics for further research.

Finally, we present an agenda for research that supports the development and evaluation of a prototype VE system for combat leader training and mission rehearsal. The topics in this plan reflects areas where there are uncertainties about task requirements, or about the effects of technology limitations. In addition, the plan reflects the need for evaluation of the overall effectiveness of instructional strategies that incorporate VE technology.

Squad Leader Tasks for Training and Mission Rehearsal

Our main source of information on relevant tasks for training or practicing using VE technology was an analysis performed by Jacobs, Crooks, Crooks, Colburn, Fraser, Gorman, Furness, & Tice (1994). That analysis described the subtask standards and specific activities included in several relevant ARTEP mission training plans for infantry and Special Forces units. The analysis also provided information about the capabilities required to perform these activities using VE technology. When necessary, we supplemented the information found in the Jacobs, et al. (1994) report with information from the following source documents:

Mission Training Plan for Infantry Rifle Platoon and Squad (ARTEP 7-8-MTP);
Battle Drills for Infantry Rifle Platoon and Squad (ARTEP 7-8-DRILL);
Mission Training Plan for the Special Forces Company: Special Reconnaissance (ARTEP 31-807-31-MTP);
Mission Training Plan for the Special Forces Company: Direct Action (ARTEP 31-807-32-MTP).

The previous analyses indicated which tasks and subtasks included activities that required voice and gesture recognition. Of the 252 activities identified by Jacobs, et al., the following five were directly relevant to recognition technologies: Give verbal orders, use password, call in preplanned fire requests, operate radio or telephone, and give hand and arm signals. Although these activities are a small proportion of the total number of activities, they are involved in a substantial proportion of the tasks that were analyzed, as shown in Table 1. This result indicates that voice and gesture recognition are an important requirement for any system using VE technology to train dismounted infantry tasks.

Table 1

Number and Percentage of Tasks that Require Each Voice or Gestural Activity (data from Jacobs, et al., 1994)

Activity	Number of Tasks	Percentage of Tasks
Give verbal orders	67	100
Use password	12	18
Call in preplanned fire requests	11	16
Operate radio or telephone	41	61
Give hand and arm signals	21	31

The requirements for CGF cannot be determined simply by examining the activities that are performed, because the activities of different squad members are not distinguished in the analysis. That is, some of the activities identified for a task are performed only by the squad

leader, and would not need to be represented by the CGF. Other activities are performed by squad members, and would need to be simulated. In addition, certain activities occur at the beginning or end of a task, and may be included or not in the simulation, depending on the training focus and the technological capabilities of the system.

Both the analysis of Jacobs, et al. (1994) and the documentation on which it was based describe several levels of detail, from missions to tasks, subtasks, and activities. Missions and tasks require a wide variety of activities, representation of at least a platoon of troops, and considerable opposing forces (OPFOR). The capability to provide for this variety and mass of troops is beyond the scope of a trainer for small unit combat leaders, and probably beyond the capability of current technology, as well. Activities, on the other hand, are often not meaningful behaviors, and consequently lack sufficient context to form the basis of training. Those that can be trained independent of context, such as hand and arm signals, can be trained easily and cheaply without technology in a classroom setting. Subtasks, either from Mission Training Plans or Battle Drills, offer a reasonable level of detail for training, and they often specify actions for single squads. Other candidates for VE training may be developed by recombining activities into especially designed training scenarios to take maximal advantage of current or projected technology capabilities. We used this latter strategy to develop representative scenarios as candidates for training using VE technology.

Using the existing task analyses as a starting point, we synthesized activities into training scenarios at three levels of difficulty. We made several considerations in developing candidate scenarios for VE training. First, the scenarios were designed for training of small unit leaders, that is, squad or fire team leaders, or leaders of Special Forces Detachments. Since the ARTEPs are focused at the platoon level, some of the tasks included in them were not appropriate at the squad level. We were also concerned with dismounted forces and focused on the tasks they perform. We stressed activities that would involve the technologies that are most relevant to this study, namely, voice and gesture recognition, and CGF. With regard to voice recognition, we limited our consideration to tasks that require speech that was structured, rather than tasks that required unstructured conversations between a combat leader and some other individual. As our summary of capabilities of voice recognition technology indicates, this level of language understanding is expected to be beyond the state-of-the-art for the next several years.

We considered training applications in all types of institutional and unit settings. However, we did not consider the initial training of how to give voice or gestural commands, because they are easy to train without any technology. The simplest scenarios we considered allow soldiers to practice giving commands in a wide variety of situations. The most complex involved larger missions such as entering and clearing a building, or conducting an assault.

We developed brief descriptions of nine scenarios at basic, intermediate, and advanced skill levels. The levels correspond roughly to technology requirements; that is, the advanced scenarios tend to require more advanced technology than more basic scenarios. These scenarios are not exhaustive. They are meant to represent the types of activities that could be trained in VE. They are not described in sufficient detail to implement them; rather, they are described so that we may understand the types of gesture recognition, speech recognition, and CGF

technology that they would require, and assess the training need that they would fill. The ability to train these tasks in VE may also require advances in other VE technologies, such as visual display technology and locomotion simulators.

The scenarios are described briefly in the following subsections. More complete descriptions are given in Appendix A.

Basic-level Scenarios

The goal in developing basic scenarios was to produce the simplest tasks for which we thought that there would be some benefit for training using VE technology. These scenarios allow the trainee to practice relevant activities in a variety of tactical situations. The technology required needs to be able to recognize trainee actions and to respond appropriately. Thus, although the benefit of implementing these scenarios is modest, the cost is also limited.

Control squad formations and movement. The purpose of this task is to give the squad leader practice in giving arm and hand signals, recognizing the conditions under which such signals are required, and showing the effects of arm and hand signals. When performing this task, the training participant is told that he will be given verbal administrative instructions during the exercise to control the movement of the squad. He must move with the squad while it is moving and maintain his proper position in the squad based on the formation. All commands must be given as arm and hand signals.

The squad is in open terrain where the squad leader can see all members of the squad. Instructions on what activities to perform are given over a headset. Directions are administrative and specific but are synchronized with what the participant is seeing. Each action or activity is 'joined' with the other activities but they are not necessarily related nor is there a requirement to be tactically realistic.

Issue fire commands. The purpose of this task is to give the squad leader practice in recognizing and organizing the organic weapons assets, recognizing and assessing the enemy situation, and controlling organic fires through voice commands. The training participant is placed in a squad in a fixed position (either a support by fire or a defensive position is easiest although some assault formations are a possibility). Actual or potential enemy locations are presented. The participant is told he must give verbal fire commands. There are commonly six elements to an initial fire command. There are also subsequent fire commands which can change the elements of the initial commands, and a cease fire or end of mission command.

Collect and report information. The purpose of this task is to give the participant practice in assessing and identifying situations accurately, organizing observations into a report format, and reporting information and communicating on a tactical radio. The participant is placed in a tactical or semi-tactical situation (e.g., observation point, check point, watch tower) with a standard or a simulated radio. He is provided a call sign and is told to report certain specified information. A variety of situations should be available including transporting or varying the presentation of the stimulus.

The participant's role is primarily passive; the emphasis is on the observing and reporting rather than the tactical response. Not all situational presentations should be obvious (i.e., an enemy soldier); some should be of "neutral" situations that the participant is expected to assess. The response by the receiver, which could include instructions like requests for clarification, more information, or continued updates, should also be incorporated into the scenario.

Intermediate-level Scenarios

The goal of the intermediate scenarios was to integrate verbal commands or hand-and-arm signals into tactically meaningful activities. These scenarios require the small unit leader to react to the situation, rather than just to verbal instructions. They require both the squad leader and other simulated squad members to perform a wider variety of activities than the basic scenarios do. However, they still do not represent complete tasks as found in ARTEP MTPs.

Conduct a dismounted patrol. The focus of the training is on a soldier who is the squad leader of a dismounted standard infantry squad on a patrol mission. A series of events, continuous but separable, are set to occur by controlling the stimulus in the form of the terrain and cover conditions, enemy, and directions given to the squad leader. In the first event, the squad leader controls formation, direction, distance, speed, and orientation of movement, as required by the terrain and actions of the squad. Movement is over moderately open terrain and there is minimal risk of contact. In the second event, contact is imminent. The squad must move by bounds, which is still relatively easy, although it is more difficult than the first event. In the third event, enemy contact is made. The squad leader controls formation, direction, distance, speed, interval, positioning, overwatch of movement. He reacts to enemy fire by controlling fire and maneuver of squad through arm and hand signals and voice commands. He reports the situation, requests and adjusts mortar fire.

Call for and adjust fire. The purpose of this task is to give the participant practice in assessing and calling for mortar or artillery fire in a tactical situation, under time pressure. The participant is a squad leader in a prepared or hasty defensive position. He will be performing as an indirect fire observer (FO). He has been told he has direct support (either artillery or mortars) and a radio, and has established communications with the fire direction center (FDC). A threat is presented that is appropriate to indirect fire support.

The FO must locate the target by one of three common methods (grid, polar, shift from known point). He must determine the location of the target and his direction to the target and must give the FDC his location. The initial call for fire requires the following elements, of which only the first four are standard: Observer identification, warning order, target location, description of target, method of engagement, method of control, and authentication.

Adjustment of fire is more complex. An initial round is fired. The FO is trying to both bring the round on line (right/left) and on range (over/short) of the target. To do this, he first must sense where the round landed in relation to the target, apply some rules of geometry, and issue a correction. Adjustments are continued with a single round until the round lands within 25

or 50 meters (depending on if it is mortar or artillery) of the desired target point. Note that this requires a high resolution visual display in addition to the technologies under investigation.

Set up and occupy hasty defensive positions. The training participant is a squad leader with a standardized or reinforced infantry squad. He is given an orientation (on a map or on the "ground") to his area and told to set up a hasty defense of a strong point or to establish a perimeter defense. He is given a general area to set up in. He must pick the exact place to defend and position his squad. He must position his squad automatic weapon (SAW) in the most likely enemy avenues of approach and position his grenadiers to cover dead space. He must provide for 360-degree defense and for overlapping fires. He should position Claymores or obstacles in areas he cannot cover. He must position observation points and provide for communication or withdrawal. He must provide for alternate and or supplementary individual fighting positions. All positions must provide for cover and concealment. He must plan for and occupy routes of withdrawal from the position and provide for rally points or supplemental squad defensive positions.

Advanced-level Scenarios

The advanced scenarios provide the greatest challenge to the squad leader, and to the simulation technology, as well. In addition, these tasks require greater coordination among the members of the unit. Consequently, simulating these tasks will require sophisticated CGF capabilities.

Enter and clear a building. The purpose of this task is to systematically enter, search, and clear a building, destroying all enemy, as part of combat in urban terrain. The training participant is the squad leader (possibly with other members of his squad). The building should be at least two stories, with a basement. The squad leader must establish the outside force and the assault force. He must select the entry point, which should be the highest point and avoid obvious entry points like doorways. Ropes, grapples, or rappelling may be required. His force must clear the entry. Inside, they organize into support teams and assault teams. Each hallway and each room must be cleared systematically. Participants must use cooked off grenades and automatic weapons fire in every room. They must employ a variety of methods in entering rooms to avoid a pattern. They must check for, discover, and disarm booby traps. They must clear obstacles. They must keep constant track of each other through voice alerts and announce all entries and exits from rooms and hiding places.

Conduct a point ambush. The purpose of this task is to select a location to ambush enemy forces, avoid detection, provide early warning, position forces, execute the ambush, destroy all enemy, and escape from the area rapidly and without casualties. The training participant is the squad leader. He is oriented to a particular location and told to prepare a point ambush. He is given the expected ambush target (dismounted troops and numbers, vehicles) and may be supplemented with special weapons or munitions (mines, anti-tank, machine-guns). He must select the site for the ambush and identify the kill zone limits. He must establish flank security and provide for early detection. He must set up mines and automatic weapons and grenades to cover the kill zone. He must provide for total concealment. He must position personnel and

designate fields of fire to cover the kill zone. He should provide for an assault force. He must institute control measures to control opening, shifting, lifting and cease fires. He must position individuals and institute control measures to avoid fratricide. He must execute the ambush to maximize the kill. He must withdraw his force rapidly and meet at a preselected point. He must minimize friendly casualties.

Conduct an assault. The purpose of this task is to practice organization, control, and conduct of dismounted assault on hasty and fortified enemy positions. An identified enemy position appropriate for a squad objective is given to the training participant, who is the squad leader. He is located in an attack position short of the objective. He must organize his assault force and his covering force and pick positions for both. He must plan for employment of indirect fires and smoke. He must provide for the lifting and shifting of organic fires and indirect fires. He executes the assault, controlling both the assault forces and the supporting forces. Fratricide is a concern and should be a measurable item. Since the squad would normally conduct an assault as part of a platoon, coordination with other units is also important.

Virtual Environment Technologies

Creating a VE and presenting it to the combat leader requires many component technologies, including the terrain data base, visual display system, direction of gaze control, movement control, and others. The goals of this section are to describe current capabilities in the three areas; voice recognition, gesture recognition, and CGF; and to predict the future capabilities, both in the near term (two years in the future) and in the somewhat more distant future (five years).

Predicting future technological advances has several difficulties. One of the major difficulties is that the extent of progress in technological capabilities depends on what resources are allocated to technological development, which, in turn, depends on the size of the market for the specific technology. VE technology has many civilian applications which will tend to control the development of these technologies. Our predictions are conditional upon the existence of sufficient levels of funding for technology development. Uncertainty regarding the level of funding is one source of error in our predictions.

We used several sources of information to assess the capabilities of VE technologies. We obtained background information on capabilities from a recent review of VE capabilities conducted by the National Research Council (Durlach & Mavor, 1995). The information in this review was supplemented by other reports and publications, interviews with researchers in the field, and searches of documentation available on the internet. The following discussion describes the current capabilities and predicts future capabilities in each of the three technology areas, based on this information.

Voice Recognition

Two aspects of voice recognition are important to the applications of VE technology to train combat leaders. The first aspect is identification of the actual words that are spoken. The last several years have seen substantial advances in the capabilities of systems to recognize spoken words, as described below. The second aspect is understanding the meaning of the sentences, so that an appropriate response may be made. The difficulty of understanding speech in this sense depends on the extent to which speech is structured or constrained. In general, language can be understood only if both the domain of discourse and the grammar are constrained.

We discuss three aspects of voice recognition technology, taken from Durlach and Mavor (1995).

- Trained vs. speaker independent. A speaker-independent system can function for a variety of speakers. In a trained or speaker-dependent system, each speaker must pronounce several words to train the system to recognize his or her voice.
- Isolated words vs. continuous speech. Voice recognizers that recognize isolated words require pauses of 100-250 ms between words or commands. Continuous speech

recognizers can recognize words in a more natural speech context, although they still require that the words be carefully pronounced and clearly stressed.

- Vocabulary size. Vocabulary size may vary from 2 to 50,000 words (Durlach & Mavor, 1995).

Other factors may affect recognition performance. For example, variation in the pronunciation of words by the same individual or unclear pronunciation can decrease the recognition accuracy of the system. In addition, the performance of the system may be sensitive to background noise. These issues are of some concern in using voice recognition technology for combat simulation, where there may be substantial background noises, and pronunciation may be affected by the stress of the simulated mission.

Word recognition capabilities. Recent improvements in speech modeling techniques and the increased power of computer workstations have produced dramatic improvements in the capability of speech recognition systems over the last several years (e.g., Nejib, 1995). These improvements have led Durlach and Mavor (1995) to state that:

> high-accuracy, real-time, speaker-independent, continuous speech recognition, for medium-sized vocabularies (few thousand words), is now possible in software on off-the-shelf workstations. (p. 234)

The described capability is clearly available in research systems, which are approaching 99 percent recognition accuracy with continuous speech for a vocabulary of at least 1000 words (e.g., Levin, Glickman, Qu, Lavie, Rose, Ess-Dykema, & Waibel, 1995; Spoken Language Systems Group, 1995; Suhm, Geutner, Lavie, Mayfield, McNair, Rogina, Schultz, Sloboda, Ward, Woszczyna, & Waibel, 1995; Tummala, Seneff, Paul, Weinstein, & Yang, 1995). Further advances are increasing the vocabulary by an order of magnitude. For example, a recent system developed by Pallett et al. (cited by Durlach & Mavor, 1995) has achieved an 11 percent word-recognition error rate using a large-vocabulary, multi-speaker, speech data base. These capabilities will become more affordable as memory and speed of computer workstations increase, and as recognition algorithms are refined and implemented in silicon, rather than software.

The capabilities of commercial systems for speech recognition, though less than those for research systems, are still substantial. Current speaker-independent systems can obtain a recognition accuracy between 85 and 90 percent with a vocabulary of about 1000 words. Trained systems can approach 95 percent accuracy for around 500 words. The choice between continuous and discrete systems depends on the performance required and the cost constraints.

Understanding speech. Understanding language is a much harder problem than merely recognizing spoken words (Abel, Reece, & Smith, 1995). Consequently, capabilities for language understanding are much less advanced than the capabilities for word recognition. In general, the requirements for language understanding are a limited domain and simple grammar.

Summary of capabilities. Current and projected capabilities in voice recognition are summarized in Table 2. Currently, both trained, isolated-word systems and speaker-independent, continuous systems can provide relatively high recognition accuracy with a vocabulary of several hundred to a thousand words. Greater accuracy is possible with trained systems than with speaker-independent systems. The choice between technologies must be based on performance requirements, number of commands to be recognized, cost, and growth potential. In the near future, speaker-independent, continuous systems will be the only reasonable choice. They will have good performance, moderate vocabularies, and reasonable cost. Large-vocabulary systems, such as the systems that are currently being developed by research institutions, should be available in the longer-term.

Table 2

Summary of Capabilities of Voice Recognition Technology

Factor	1996	1998	2001
Trained vs. Speaker-Independent	Trained has the lower error rate	Speaker independent	Speaker independent
Isolated Words vs. Continuous Speech	Either appropriate, depending on cost and performance requirement	Continuous speech with moderate vocabulary size	Continuous speech with large vocabulary size
Vocabulary Size	Hundreds to a few thousand	A few thousand	Ten thousand or more
Noise Tolerance	Requires low noise	Requires low noise	Higher noise tolerance

The capabilities for language understanding are much less advanced, and the future of these capabilities is more difficult to predict. For the foreseeable future, it appears that voice recognition will be most applicable to situations involving simple, structured commands, or where the voice interactions are otherwise constrained.

Gesture Recognition

A gesture recognition system must perform two component processes. First, it is necessary to locate and track the parts of the body giving a gesture in space. Then, it is necessary to interpret the movements as a particular command or other communication. There are many technologies for position tracking, each with different strengths and weaknesses. Furthermore, these technologies may be combined to form hybrids that can provide the advantages of their component technologies. The problem of interpreting movements has received less attention than the parallel problem in voice recognition. Nevertheless, a moderate amount of progress has been made in the ability to recognize certain gestures.

Position tracking. Durlach & Mavor (1995) have enumerated the following technologies for performing position tracking.

- Mechanical trackers. This category includes body-based linkages (e.g., exoskeletons), which attach to the body, and ground-based linkages, which attach to a fixed location on the ground. These trackers can be inexpensive and fairly accurate. Although ground-based linkages are not appropriate for sensing hand and arm signals, body-based trackers have some potential, if problems with fit, measurement, alignment, and calibration can be solved. These trackers have the disadvantage of limiting the mobility and comfort of the user, who either must be attached to a fixed mechanical device, or must be attached to potentially encumbering cabling.

- Magnetic trackers. Magnetic trackers are popular, because of their low cost, reasonable accuracy, and convenience. Though they do not restrict the user as much as mechanical trackers, they still require the user to be tethered to the system with cabling. These trackers are limited by the latency of their response, and because their accuracy may be reduced due to interference from extraneous magnetic fields from other equipment. Sensors that use DC magnetic fields are less prone to interference than those that employ AC magnetic fields.

- Passive stereo vision systems. These systems use one or more cameras to provide input to the location tracking system. According to the National Research Council review, these sensors are not likely to be useful to VE applications in the near term. They may prove useful in the long term, however, after the technology becomes more developed.

- Optical marker systems. In this method, markers are placed on certain critical locations on the body, which are then tracked with cameras. Problem with this technology include the potential for obscured visibility, if markers are blocked due to the position of the body.

- Structured light systems and laser radar. These two technologies are currently available in research systems. Although they do not have sufficient accuracy for practical applications, their accuracy is improving, and they may prove to be viable alternatives in the future.

- Laser interferometers. These methods are accurate, but expensive. They also measure relative distance, rather than absolute distance. Finally, there is difficulty in tracking multiple limbs, as well as potential visibility problems.

- Acoustic trackers. Acoustic trackers are the basis of inexpensive sensor devices, including the Mattel Power Glove. They are inexpensive, but limited in accuracy, speed, and range because of interference, echoes, and atmospheric attenuation. Furthermore, they require that the user be tethered to the system with cables. Tracking multiple markers will take further advancements, such as using multiple frequencies.

- Inertial tracking. Inertial sensors are currently too big and too expensive. However, these sensors could be combined with another technology, such as acoustic trackers, to produce a cost-effective tracking technology. They avoid the need for tethering by using a radio-frequency (RF) link to the computer.

Gesture interpretation. There are a variety of methods for gesture interpretation, including artificial neural networks, fuzzy sets, template matching, trajectory matching, and petri nets (Searles, Smith, Baratoff, & Bohmueller, 1993; Abel, Reece, & Smith, 1995). Some methods are based on the same general procedures used for voice recognition, such as voice recognition, but they are tailored to the specific requirements of gesture recognition.

There is some research that has determined the recognition rates for military hand and arm gestures. Searles et al. (1993) obtained correct recognition rates of 96% over a set of seven gestures using a trajectory matching algorithm. This value was considerably better than the 82% that was obtained overall using a template matching approach. Recognition was better for static gestures that involve a single arm position, than for dynamic gestures that involve two or more positions. The gestures were given in standard conditions in which the position of the individual making the gestures relative to the sensors was controlled. In a tactical scenario, gestures would be given in a variety of positions and orientations. Nevertheless, the results are very promising regarding the potential for gesture identification technology.

Summary of capabilities. Current and projected capabilities in gesture recognition are summarized in Table 3. Currently, there are multiple sensor technologies, each with its own strengths and weaknesses. Many of these technologies have not matured to the point where they are available in off-the-shelf systems. It is possible that one of these new technologies will prove to be a cost-effective method for position tracking. In the short term, improvements in performance may be made using hybrids of existing systems, such as magnetic, acoustic, and inertial sensors.

Table 3

Summary of Capabilities of Gesture Recognition Technology

Factor	1996	1998	2001
Sensor Technology	Multiple technologies	Hybrid systems	Mature technologies
Recognition Procedures	Better performance for static gestures	Static and dynamic gestures	Correlation with voice

Techniques for recognizing gestures from sensor inputs have not been thoroughly evaluated, especially in the widely varied conditions that would be expected to occur in tactical simulations. Recognition accuracy is better for simpler, static gestures than for more complicated, dynamic gestures, but the indications are that performance should improve for all gestures in the near

future. Correlation with other sensory inputs and recognition of more subtle, informal gestures will require additional time and effort.

Computer-Generated Forces

CGF represent all squad members and OPFOR in the scenarios. They must respond realistically to spoken commands and gestures given by the leader. The behavior of the CGF must also be appropriate for the mission goals, terrain, and enemy strength and location.

In addition, the actions of the CGF must be presented realistically to the unit leader. The leader must be able to verify that the unit members are in the desired locations, and are using the proper movement techniques. In addition, the leader needs to receive communications from unit members, either verbally or through hand-and-arm gestures.

CGF for VE training of small unit leaders will require the combination of two kinds of technologies. The first technology represents the information processing capabilities of unit members. This technology will control how unit members select their movement paths and use their weapons. Existing methods have been developed to control friendly and enemy units in distributed interactive simulations. The second technology visually portrays the activities of unit members to the leader, allowing for two-way communication between the leader and the simulated unit. Existing capabilities have been developed by the entertainment industry and human factors researchers. The following two subsections give a brief description of the capabilities in these two areas.

Generally speaking, there will be differences in the requirements for CGF representing OPFOR and those representing squad members. OPFOR CGF will need to demonstrate tactically correct unit behavior that is consistent with the training objectives of the scenario. Squad member CGF, on the other hand, need to demonstrate individual behavior that is consistent with the direct commands of the squad leader, each individual's role in the squad, required tactical behaviors, and the capability to interpret and act upon the squad leader's orders realistically.

Information processing. One of the major features of training with DIS has been the development and use of CGF to control both enemy and friendly forces with minimal intervention by a human controller. The ModSAF software that controls simulated entities in SIMNET allows the controller to specify mission goals and activities for CGF (Loral, 1995). The software then can plan the specific movements of the entities to reach goals, avoid obstacles, and follow roads as required. There is some capability to avoid moving, as well as fixed obstacles, but meeting multiple constraints has been a difficulty for the model (Smith, 1994). ModSAF also has the capability to search terrain to find covered and concealed positions (Longtin, 1994) for armored vehicles.

Recent research has been focused on producing greater autonomy of CGF. It has also attempted to increase the capability related to dismounted infantry. Some of the problems that are being addressed include improved terrain reasoning, movement control, situation assessment, tactics, and cooperation (Reece, 1994).

Presentation of CGF. Representing human figures is much more difficult than representing vehicles, because there are many more degrees of freedom for human movement (Reece, 1994). Currently, ModSAF software has limited ability to represent dismounted infantry. For example, ModSAF allows dismounted soldiers to be in one of three postures, standing, kneeling, or prone (Loral, 1995). Thus, in this system, simulated soldiers do not move realistically and cannot communicate by gestures to the unit leader. A simplistic representation of a human recently demonstrated as part of the Dismounted Soldier Simulation demonstration used 1600 polygons to represent the movements of a single figure. This number is considerably greater than the around 200 to 300 polygons that are required to represent a ground vehicle. This demonstration can represent many body movements. However, if individual finger movements must be presented, then additional polygons would be required. As the number of individuals within the scene increases, so will the image generation requirements to allow appropriate recognition of gestures.

Advancements in entertainment software and human factors have produced more realistic graphical representations of human figures. One such system, developed at the University of Pennsylvania, is the *Jack* system (NASA, 1992). *Jack* includes a detailed three-dimensional human model, including 69 segments, 68 joints, and 121 degrees of freedom. This system can present a realistic human representation with reasonable freedom of movement. The growth of these capabilities is currently accelerating as the interest in "virtual reality" increases and the capabilities become more affordable. Systems such as the Integrated Unit Simulation System (IUSS) are being used to analyze current soldier system performance. Refinements in this technology could lead to extremely realistic human models in the future.

Summary of capabilities. Current and projected capabilities in CGF are summarized in Table 4. The table shows that future advancements will produce greater levels of tactical knowledge and team coordination, as well as improved visual presentation. The near future should produce CGF with the kind of presentation capability that is currently available in the *Jack* system. Further advancements will produce increased realism at a lower cost as computing capabilities continue to advance.

Table 4

Summary of Capabilities of Computer-Generated Forces

Factor	1996	1998	2001
Information Processing	Fixed obstacle avoidance, predefined alternatives	Team coordination, limited terrain reasoning	Analyze terrain and enemy
Gesture Presentation	Limited hand and arm gestures	Detailed movement at close range	Freedom of movement
Human Representation	Cartoonish, exaggerated expressions	More realistic and complex	Realism at lower cost

Functional Requirements For Scenarios

The previous two sections of this report described nine illustrative scenarios that could form the basis for VE training for small unit leaders and outlined the current and future capabilities of VE technologies for voice recognition, gesture recognition, and CGF. This section describes the functional requirements of the scenarios, in terms of the characteristics that were used to describe the technologies. As a result of this analysis, we will make a preliminary assessment of which scenarios can be supported by current and projected future technology.

The main result of this analysis is shown in three tables that describe the assessed requirements for voice recognition, gesture recognition, and CGF. The tables estimate the number of words of gestures that the system must recognize, whether words are spoken in isolation or as a part of continuous speech, whether gestures are static or dynamic, the kinds of CGF activities that must be performed, the CGF activities that must be presented to the soldier, and the other characteristics that were used to assess the capabilities of the VE technologies. The assessments reflect the information shown in the scenario descriptions (see Appendix A), as well as the knowledge of an infantry subject matter expert (SME) on the project staff who reviewed them and corrected inaccuracies.

We used several assumptions to guide the assessment of the requirements of the scenarios for some factors. For example, we specified speaker independent voice recognition as a requirement when there were more than 50 words in the vocabulary, or when the set of words could not be specified in advance. This assumption simply reflects the opinion that training a voice recognition system more than 50 words would present an unacceptable inconvenience to the trainee. Similarly, for the information processing requirements for the CGF, we assumed that movements of simulated soldiers could be predefined if the starting positions for the scenario were fixed and there were no enemy contact.

The factors used to rate the functional requirements for gesture recognition are different from those used in the technology description. Specifically, we assessed requirements for gesture recognition according to whether the required gestures are static or dynamic, the number of gestures that could be used in a scenario, and the parts of the body that are used to perform the gestures. These factors were easy to assess based on the scenario descriptions, while the factors that were used to describe the technology (sensor technology and recognition procedures) did not seem to characterize the requirements adequately. In particular, it was very difficult in most cases to state that a given scenario required a specific sensor technology. In most cases, several technologies could be adequate, although some could be eliminated, and others would require further research to develop adequate capabilities.

There is considerable uncertainty in the requirements, which reflects lack of research knowledge of the links between technological capabilities and training effectiveness. For example, the relationship between recognition error rates and training effectiveness is unknown. Consequently, we can neither derive requirements for recognition accuracy, nor make assumptions regarding the accuracy rate. In a later section of the report, we recommend this topic as a candidate for future research.

Basic-level Scenarios

The functional requirements for the basic-level scenarios are shown in Table 5. Comparing these requirements to the capabilities indicates that most of the requirements of the basic scenarios can be met with current technology, with some caveats. Because of its larger vocabulary, and because of the longer messages that must be communicated, collecting and reporting information was assessed to require speaker-independent recognition of continuous speech. Isolated word recognition was judged to be inadequate because the reports are too long to give on a word-by-word basis. The required level of performance is within the limits of existing technology if the recognition rate is adequate.

The scenario *Control of Squad Formation and Movement* was designed to require communication with gestures only. Consequently, it does not require voice recognition capabilities. Both the head and eyes are relevant to communication using gestures, in addition to the arms and hands. The direction the head is turned gives a focus and direction to the gesture and may indicate the intended recipient of the gesture.

Intermediate-level Scenarios

The functional requirements for the intermediate-level scenarios are shown in Table 6. The scenarios that require gestures generally require the same set of tactical gestures that are used in the basic scenarios, although they may be given in a wider variety of situations. Also, there may be several sequences of gestures that can accomplish a tactical goal in the intermediate scenarios, perhaps requiring somewhat more sophistication in the gesture recognition algorithms. In summary, however, the gesture requirement for these scenarios are not much different than those for the basic scenarios.

The variation in requirements for speech recognition is similar to that for the basic scenarios. The scenario *Conduct a Dismounted Patrol* was designed to have minimal verbal requirements that are well within current capabilities. The scenario *Call For and Adjust Fire* requires a fairly structured exchange of information, which also seems within existing capabilities. The speech-recognition requirements for *Set up and Occupy a Hasty Defensive Position* are greater than the other scenarios, because it requires the leader to issue fairly unstructured commands. The required speech-recognition capability may be beyond current capabilities, but they will almost surely be within the capabilities in the near future.

The requirement for team coordination, and the requirement that two of the scenarios place on the CGF system are the major obstacles to training these scenarios using VE technology. The scenario *Conduct a Dismounted Patrol* requires a level of team coordination that is consistent with the projected capability in the near-term future. Setting up a hasty defensive position requires even greater team coordination, consistent with the longer-term projection.

Table 5

Functional Requirements of Basic Scenarios

Factor	Control Squad Formations and Movement	Issue Fire Commands	Collect And Report Information
Voice Recognition			
Trained vs. Speaker Independent	NA	Trained may be adequate.	Speaker independent required.
Isolated Words vs. Continuous Speech	NA	Isolated words may be adequate.	Continuous speech required.
Vocabulary Size	NA	25-30	150-300 words
Noise Tolerance	NA	High, if weapons noises are simulated	Low, communication in relatively quiet environment.
Gesture Recognition			
Static vs. Dynamic	Full range of gestures required.	NA	NA
Gesture Vocabulary Size	25-30	NA	NA
Relevant Body Parts	Hands, arms, head, and eyes	NA	NA
Computer-Generated Forces			
Information Processing	Most alternatives can be predefined, since scenarios are brief. CGF may need to make errors to be corrected by leader	Must display realistic methods, e.g., automatic engagement of emerging point targets, dispersion of fires.	Friendly, enemy, and neutral objects in preprogrammed activities
Gesture Presentation	Response gestures	Minimal	None
Human Representation	Leader must discern CGF movement methods.	Minimal	May need to represent certain threatening activities

Table 6

Functional Requirements of Intermediate Scenarios

Factor	Conduct a Dismounted Patrol	Call For and Adjust Fire	Set up and Occupy Hasty Defensive Positions
Voice Recognition			
Trained vs. Speaker Independent	Required at enemy contact. Trained may be adequate.	Speaker independent required.	Speaker independent required.
Isolated Words vs. Continuous Speech	Isolated words may be adequate.	Isolated words may be adequate.	Continuous speech required.
Vocabulary Size	25-30	100-150	Several hundred words
Noise Tolerance	High, if weapons noises are simulated	High, if weapons noises are simulated	Moderate noise from soldier activities
Gesture Recognition			
Static vs. Dynamic	Full range of gestures required	NA	Full range of gestures required
Gesture Vocabulary Size	25-30	NA	25-30
Relevant Body Parts	Hands, arms, head, and eyes	NA	Hands, arms, head, and eyes
Computer-Generated Forces			
Information Processing	Requires coordinated movement over terrain using proper methods. CGF must take appropriate actions at contact.	Must display realistic methods.	CGF must perform variety of activities and respond to each other and to leader commands. Needs intelligent OPFOR.
Gesture Presentation	CGF must give gestures in response to leader commands.	Minimal	Wide variety of movements presented at fairly close range
Human Representation	Leader must discern tactical formations and weapon status.	Minimal	Realistic presentation beneficial because of variety of activities.

Advanced-level Scenarios

The functional requirements for the advanced-level scenarios are shown in Table 7. All three of the advanced scenarios are beyond the current capabilities of VE technology. All require speaker-independent recognition of continuous speech, with a moderate vocabulary. This should be within the capability of speech recognition systems in the near future, if the domain of discourse is sufficiently constrained so that the meaning can be interpreted correctly. However, in all three technology areas, there is a requirement for considerable flexibility. Spoken messages cannot be specified in advance. Soldiers may use informal gestures that rely on common understanding rather than formal definitions, especially in the scenario *Enter and Clear a Building*. Finally, the scenarios require a high level of team coordination, autonomous actions of team members, and a realistic and detailed representation of team-member actions.

Table 7

Functional Requirements of Advanced Scenarios

Factor	Enter and Clear a Building	Conduct a Point Ambush	Conduct an Assault
Voice Recognition			
Trained vs. Speaker Independent	Speaker independent required.	Speaker independent required.	Speaker independent required.
Isolated Words vs. Continuous Speech	Continuous speech required.	Continuous speech required.	Continuous speech required.
Vocabulary Size	Words cannot be specified in advance. May need a few thousand words.	Words cannot be specified in advance. May need a few thousand words.	Words cannot be specified in advance. May need a few thousand words.
Noise Tolerance	High, if weapons noises are simulated.	Low, until enemy is engaged; then high.	High, if weapons noises are simulated.
Gesture Recognition			
Static vs. Dynamic	Both required	Both required	Both required
Gesture Vocabulary Size	Informal gestures possible	10 or fewer	25-30
Relevant Body Parts	Hands, arms, head, and eyes	Hands, arms, head, and eyes	Hands, arms, head, and eyes
Computer-Generated Forces			
Information Processing	High level of team coordination. Autonomous actions of team members.	Flexibility needed to execute leader's plan. Moderate team coordination.	High team coordination and terrain reasoning required
Gesture Presentation	Freedom of movement at limited range	Freedom of movement at a limited range	Freedom of movement
Human Representation	Realistic	Realistic	Realistic

Developing a Research Agenda

The previous discussion identified several uncertainties regarding the capabilities of VE technologies to represent critical aspects of combat leader tasks. In addition, the effects of technological limitations on performance and training is not known. This section begins by enumerating a set of research questions that represent critical gaps in our knowledge about the effectiveness of VE technology for performing and training combat leader tasks. It then outlines a program of research to address these uncertainties.

Research Questions

The following nine questions address a variety of issues, from the performance of specific technology options to the overall effectiveness and efficiency of VE training for combat unit leaders. After stating each question, we briefly elaborate on some of the issues involved.

- *What are the error rates for voice and gesture recognition in tactically realistic scenarios?*

The tactical scenarios that are used for training present conditions that are considerably different from those in which a manufacturer evaluates its voice or gesture recognition device. Environmental conditions and the stress of the simulated combat may affect how gestures verbal commands are given, and consequently, the accuracy with which they are recognized. It is important, therefore, to assess the performance of the recognition devices under realistic conditions.

- *How do errors in recognizing voice or gesture commands affect performance or training effectiveness?*

To some extent, recognition errors are a natural reflection of performance in the field, where commands may also be misunderstood or misinterpreted. However, if the error rate is too high, then performance will suffer, because the combat leader will need to repeat commands when they are not understood, or the system CGF will perform the incorrect activities in response to a command. Recognition errors are a particular concern for training, because these errors can lead to incorrect feedback being given to the trainee. It seems likely that feedback errors would hinder training effectiveness and possibly lead to negative transfer, if they are great enough. Consequently, the effect of recognition errors on performance and training is an important question for empirical research. The results of this research could be used to set performance requirements for recognition systems used in VE combat trainers.

- *Do discrete voice recognition systems provide for an acceptable interface for issuing orders such as fire commands?*

Discrete voice recognition systems require the speaker to put a slight pause between words. While this interface may be adequate for some purposes, it is almost certainly inappropriate for

longer, more complex commands. Of particular concern here is simpler commands that have several components, such as fire commands. If these commands can be recognized by discrete voice recognition systems, then it may be possible to provide effective low-cost training using current technology. Otherwise, the development will require the additional cost and development time for continuous voice recognition systems.

- *What type of sensor technology provides the best input for gesture recognition?*

As earlier discussion indicated, there are currently several competing sensor technologies that could be used for gesture recognition. These technologies need to be evaluated to determine which technology or combination of technologies are best suited to the gestures used by the combat unit leader. Answering this question will require both analysis of the gestures to determine the sensor requirements, and empirical comparison of different technologies.

- *How are performance and training effectiveness affected by simplifications and inaccuracies in the representation and presentation of computer-generated forces?*

Leading computer forces is not the same as leading people. For example, the CGF's behavior may be somewhat stereotyped, and may show less variability than human behavior. When the leader gets back to his squad, he may be faced with unforeseen squad behavior that he does not know how to deal with. The squad leader might also learn tactics that are specific to the quirks of his CGF squad. Some of these may not transfer to the actual situation. On the other hand, the CGF squad might exhibit some unrealistic behavior which might produce situations that lead to low transfer of training. In mission rehearsal, unrealistic behavior of the CGF may cause the combat leader to lose confidence in the value of the technology and, consequently, not to use the system. The behavior of the CGF is even more important in the more advanced scenarios, because they require more complex behaviors from the simulated forces.

- *How can instructional strategies, such as augmented cues, reset, or replay, enhance training in a virtual environment?*

Simulated training allows for the application of strategies that can enhance training effectiveness and compensate for some of the deficiencies of the simulation quality. Research in the use of some of these strategies in weapon system simulators, such as flight simulators, has shown that they can be useful methods for enhancing training effectiveness. Since the tasks performed by dismounted infantry are different in many respects from those performed in weapon system simulators, we anticipate that instructional strategies would be applied differently in these tasks. It will require analytical research to determine how different instructional features would be used in VE training for combat unit leaders, as well as empirical research to evaluate the impact of these features on training effectiveness.

- *Does practice on a particular task in a virtual environment produce an improvement in performance in the field?*

The ultimate measure of effectiveness of a training system is the extent to which practice using the system leads to improved performance in actual combat situations. Transfer of training has always been difficult to assess empirically. Assessing transfer for a combat leader training presents additional problems, because field performance depends on the performance of the unit as well as its leader. Thus, the evaluation must take into account the fact that training is conducted with a simulated squad, while performance occurs with a real squad, which may perform better or worse than the simulated squad.

- *How much field training can be replaced by VE training?*

This question is concerned with the efficiency of training. The field training that is replaced by VE training can be used to justify the use of simulated training because of reduced cost. However, VE training may provide the combat unit leader with skills that are not currently trained by any methods.

- *Does performing tasks in VE lead to disorienting side effects? If so, how can they be reduced?*

Investigating side effects of immersion, including simulator sickness, has been an ongoing research concern of ARI and other agencies. This research should continue, because VE training may have a detrimental effect on trainee well being, and because side effects may have a negative impact on training effectiveness.

Steps in the Research Plan

Research and technology development are interactive activities, with each supporting the other. Early research will specify the technology requirements and identify the key performance variables. Later research will assess the performance of components and evaluate overall training effectiveness and efficiency. The first step in the plan is to develop scenarios for testing. Those scenarios will then be used to evaluate the performance of VE technologies, instructional strategies, and the overall effectiveness and efficiency of VE training.

Develop testing scenarios. The goal of this step is to develop scenarios for testing VE capabilities and assessing the effectiveness of VE systems. Scenarios should be developed to use different system capabilities. For example, the scenario *Call for and Adjust Fire* exercises the voice recognition subsystem, while the scenario *Conduct a Dismounted Patrol* exercises the gesture recognition subsystem. Tasks that require both voice and gestures should also be included. Finally, scenarios should place different levels of stress on the CGF system, by requiring different levels of coordination between unit members, or by requiring more or less autonomy for the simulated soldiers.

The scenarios outlined in this project are representative of the types of tasks that would be required. These scenarios could provide the starting point for developing testing scenarios, although some additional ones may be required. Basic scenarios are most appropriate for early

evaluations, and more advanced scenarios for later evaluations. The scenarios would be implemented in several virtual environments, including both rural and urban areas. In addition to the scenarios themselves, this effort should develop performance measures that can be applied on the system or in the field, to evaluate transfer of training.

Evaluate alternative technologies. The goal of this task is to evaluate the performance of alternative VE technologies. Especially relevant in this step is gesture recognition sensor technologies and recognition algorithms, and CGF models, but voice recognition (or other) technologies could be evaluated here, as well. In addition to the specific sensors employed, this research could address issues regarding the number of sensors, their placement, or other aspects of sensor configuration. This research would identify the strengths and weaknesses of the technologies using objective performance measures, such as recognition error rate or false recognition rate, in tactical scenarios.

Preliminary evaluation of sensor and recognition technologies can be accomplished by examining written specifications for the technology. The most attractive candidates would then be obtained and linked to the prototype VE system. Recognition algorithms would need to be modified to be compatible with the information provided by the sensors. The modifications might be extensive if very different sensor technologies were being compared, such as visual and magnetic sensors.

Evaluate training and performance issues. The goal of this research is to evaluate how the characteristics and limitations of VE technology affect the ability to train and perform tactical scenarios using the technology. Individual experiments would investigate the effects of recognition accuracy, discrete voice recognition and CGF behavior on training and performance. These are critical issues, because inaccurate recognition or oversimplified CGF models can give the trainee inappropriate feedback that might slow the progress of training. Consequently, this research would be designed to determine the maximum recognition error rate that will not detract from learning or performance. The results could then be used to set the performance requirements of recognition systems and CGF models. In addition to the specific topics described above, research in this area should continue to address measures of presence and unwanted side effects.

Some experiments could be conducted before the recognition technology is procured or developed by using humans as substitutes for the technology. For example, the effect of recognition accuracy can be simulated using a human recognizer who randomly inserts errors in to the recognition process. Similarly, simulated forces could be controlled by people who introduce specific errors to examine their effects on performance or training.

Assess instructional features. Simulated training, whether on individual weapon system training devices or DIS, provides a variety of features that are designed to enhance the effectiveness or efficiency of training. These features can set the initial conditions for a simulation, control some of the performance parameters, provide cues to the trainee, augment the feedback given to the trainee, organize information for after action reviews, and provide many other kinds of support to the instructional process. Some instructional strategies capitalize on the

ability of simulation-based training to enhance learning, while others compensate for limitations of the technology.

Research in this area would be conducted to identify, develop, and evaluate instructional features that are especially suited to the capabilities and limitations of VE technology, as well as strategies for their use. Most strategies would be analogs to those used with individual weapon system simulators and distributed interactive simulation; others would be developed specifically for VE training. The research will evaluate the effectiveness of the instructional strategies, according to the extent to which performance on the simulation system improves. This research would not investigate transfer of training.

Determine training effectiveness and transfer. The primary measure of the effectiveness of VE technology for training is the extent to which skills learned in VE transfer to performance in the field. Assessment of transfer of training requires careful experimental design and large samples of trainees. In many cases, the goal of this research is to determine how much real-world training can be replaced by a given amount of VE training, maintaining a given level of performance. Answering this question requires the use of several experimental groups that receive different amounts of VE training, and consequently have even greater cost.

Evaluating transfer of training for a device that trains small unit leaders raises several methodological and experimental design issues. One such issue involves the role of the other squad members in the evaluation. The leader is trained in VE, but overall performance depends on the activities of the entire squad. That is, training is conducted with a simulated squad, while field performance is assessed with a real squad. In order to assess transfer of training, ways must be developed to measure leader performance that are insensitive to differences in performance of other squad members.

Summary

Lack of knowledge regarding certain key questions makes it difficult to set performance requirements for VE technologies that could be used to train small unit leaders, to rehearse missions, or to evaluate system and organizational concepts. The research described in this section seeks to answer those questions by investigating the effects of technology limitations on performance, as well as the instructional strategies that can compensate for these limitations or offer other training and performance benefits. This research will help to guide the technology development process to produce the most effective system design.

Other elements of the research described in this section evaluate the performance of systems incorporating VE technology to train small unit leaders. This research assesses the improved performance resulting from VE training, and the extent to which this improvement transfers to performance in the field. The results will provide feedback to the development process, as well as rationale for implementation of the system.

References

Abel, K., Reece, D., & Smith, M.J. (1995). *Voice and gesture input for computer generated forces.* Orlando, FL: Institute for Simulation and Training.

Department of the Army (1988). *Mission training plan for the infantry rifle platoon and squad* (ARTEP 7-8-MTP). Washington, DC: Author.

Department of the Army (1990). *Battle drills for the infantry rifle platoon and squad* (ARTEP 7-8-Drill). Washington, DC: Author.

Department of the Army (1989). *Mission training plan for the special forces company: Special reconnaissance* (ARTEP 31-807-31-MTP). Washington, DC: Author.

Department of the Army (1989). *Mission training plan for the special forces company: Direct action* (ARTEP 31-807-32-MTP). Washington, DC: Author.

Durlach, N.I., & Mavor, A.S. (Eds.). (1995). *Virtual reality: Scientific and technological challenges.* Washington, DC: National Academy Press.

Jacobs, R. S., Crooks, W. H., Crooks, J. R., Colburn, E., Fraser, R. E. II, Gorman, P. F., Maden, J. L., Furness, T. A. III, & Tice, S. E. (1994). *Training dismounted soldiers in virtual environments: Task and research requirements* (Technical Report 1011). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (AD A286 311)

Levin, L., Glickman, O., Qu, Y. Lavie, A., Rose, C., Ess-Dykema, C., & Waibel, A. (1995). Using context in machine translation of spoken language. *Theoretical and Methodical Issues in Machine Translation: Vol. 1* (p. 1-1).

Longtin, M. (1994). Cover and concealment in ModSAF. *Proceedings of the fourth conference on computer generated forces and behavioral representation*, pp. 239-248. Orlando, FL: Institute for Simulation and Training.

Loral Advanced Distributed Simulation (1995). *Functional description document for ModSAF, V1.5.1* (LADS Document Number 95007 v1.0). Orlando, FL: Author

NASA (1992). Untitled summary of *Jack*. *NASA Tech Briefs*, *16* [on-line]. Available: http://ntas.techtracs.org/4b.acgl$twDisplech:Item(15;ARC-12758-1;20016307;1)

Nejib, P. (1995). *Forward Area Language Converter.* XVIII Airborne Corps, Ft. Bragg, NC.

Reece, D.A. (1994). Extending DIS for individual combatants. In *Proceedings of the 5th annual conference on AI, simulation, simulation, and planning in high autonomy systems: DIS environments.* Piscataway, NJ: IEEE Computer Society Press.

Searles, D., Smith, J., Baratoff, G., & Bohmueller, B. (1993). *Recognition of signals for combat formations and battle drills [on-line]*. Available: http://www.hitl.washington.edu/scivw/JOVE/articles/dsgbjsbb.txt

Smith, J.E. (1994). Near-term movement in ModSAF. In *Proceedings of the fourth conference on computer generated forces and behavioral representation,* pp. 249-260. Orlando, FL: Institute for Simulation and Training.

Spoken Language Systems Group (1995). *Research and Development of Multilingual Conversational Systems.* Laboratory for Computer Science, Massachusetts Institute of Technology.

Suhm, B. Geutner, P.T., Kemp, T., Lavie, A., Mayfield, L., McNair, A., Rogina, I., Schultz, T., Sloboda, T., Ward, W., Woszczyna, M., and Waibel, A. (1995). JANUS: Towards multilingual spoken language translation. *Proceedings of the ARPA Spoken Language Technology Workshop, Vol. I,* pp. 221-226. Lexington, MA: Lincoln Laboratory, MIT.

Tummala, D., Seneff, S. Paul, D., Weinstein, C., and Yang, D. (1995). CCLINC: System Architecture and Concept Demonstration of Speech-to-Speech Translation for Limited-Domain Multilingual Applications. In *Proceedings of the ARPA Spoken Language Technology Workshop.* Lexington, MA: Lincoln Laboratory, MIT.

APPENDIX A
TRAINING SCENARIO OUTLINES

Title: Control Squad Formations and Movement

Level: Basic

Purpose: To give the squad leader practice in giving arm and hand signals, recognizing the conditions under which such signals are required, and showing the effects of arm and hand signals.

Tasks and Activities: Use Visual Signals, Control Movement of a Squad, Move as part of a Squad Formation, Perform Battle Drills

Scope: The training participant is placed in a squad. He is told that he will be given verbal administrative instructions during the exercise to control the movement of the squad. He must move with the squad while it is moving and maintain his proper position in the squad based on the formation. All commands must be given as arm and hand signals.

The squad is in open terrain where the squad leader can see all members of the squad. Instructions on what activities to perform are given over a headset. Directions are administrative and specific but are synchronized with what the participant is seeing. Each action or activity is 'joined' with the other activities but they are not necessarily related nor is there a requirement to be necessarily tactically realistic.

EXAMPLES:

Situation: Squad is stationary but dispersed. Instruction: "You must form your squad into a column and move them to (location or direction). Performance Requirement: Participant signals ATTENTION, signals COLUMN FORMATION, signals MOVE OUT, takes proper squad leader's position in the formation.

Situation: Squad is moving in a column formation. Instruction: "Form your squad into a wedge." Performance Requirement: Participant signals ATTENTION, signals WEDGE, takes proper squad leader's position in the formation.

Situation: Squad is moving in a (specified) formation. One of the fire teams is at the correct interval; the other is (too close) (too spread out). Instruction: "Check your fire team's interval and take any corrective action required." Performance Requirement: Participant

recognizes incorrect fire team interval, signals (points) to that fire team leader only, issues signal for (CLOSE UP) (DISPERSE).

Situation: Squad is stationary and dispersed. Location imposed restrictions on aircraft (woods, ground obstructions, overhead wires). Instruction: "You have arrived at the LZ where your squad is to be picked up by helicopter (specify type). You must mark the LZ and bring the aircraft into your location. You have (panels) (smoke) (???) available to mark the landing point." Performance Requirement: Participant marks the landing point, identifies the location and direction of the aircraft, positions himself correctly to the landing point, gives necessary signals to put the aircraft on the landing point.

Simulation Specifics: Needs a SAFOR squad that will respond only to 'correct' (with some degree of tolerance for individual differences) arm and hand signals. May need some 'transported' or programmed terrain changes.

Notes: There are probably 25 to 30 basic individual arm and hand signals. Many are used in combination or can be situationally arranged in combinations to give almost an unlimited variety of circumstances to give 'fresh' training stimulus even to the same participant. Plus the terrain can be altered (jungle, urban) to insure perceptual if not performance variety.

Title: Issue Fire Commands

Level: Basic

Purpose: To give the squad leader practice in recognizing and organizing the organic weapons assets, recognizing and assessing the enemy situation, and controlling organic fires through voice commands.

Tasks and Activities: Identify Enemy Locations, Control Organic Fires, Issue Verbal Fire Commands

Scope: The training participant is placed in a squad in a fixed position (either a support by fire or a defensive position is easiest although some assault formations are a possibility). Actual or potential enemy locations are presented. The participant is told he must give verbal fire commands.

There are commonly six elements to an initial fire command. There are also subsequent fire commands which can change the elements of the initial commands. There is also a CEASE FIRE or END OF MISSION command. Initial fire commands and examples are:

Alert: SQUAD, TEAM ALPHA, SAW, GRENADIERS

Direction:	LEFT FRONT, WOODLINE, REFERENCE: ROAD JUNCTION - RIGHT TWO FINGERS, (tracer method)
Target Description:	TROOPS, TRUCK, COLUMN, WINDOWS, BASE OF TREES
Range:	(In meters) FOUR HUNDRED, ONE HUNDRED
Method of Fire:	Specifies who if different from Alert. Also can specify type and/or amount of ammunition. Can be omitted or an "SOP" for amount and type can be specified in the instructions. Participant can change amount/type based on the situation.
Command to fire:	FIRE, AT MY COMMAND, WATCH MY TRACER

Simulation Specifics: Need a SAFOR that is responsive to the fire commands including following the instructions for who fires, what is fired, when it fires, where it fires, so that the participants assessment and reaction to the situation can be judged. At the same time, the SAFOR has to be programmed so that it portrays realistic activities of the squad. (For example, point targets that emerge or are identified during the execution are engaged automatically by the individual riflemen; dispersion of fires over area targets are a usually a squad SOP, e.g., right, center, left coverage)

Various target arrays or presentations are required. These should be changeable by transporting or other presentation options. OPFOR should be included.

Notes: A variety of situations is intended; the content of the fire command is situational. Part of the training emphasis is reacting properly to the situation presented. Some type of assessment based on actual 'hits' should be included so the effectiveness of the fire command can be illustrated. This should include enemy who are hidden in area targets. An "advanced" table is possible where verbal fire commands are combined with visual fire commands (signals for OPEN FIRE, CEASE FIRE, SHIFT FIRE, TRAVERSE, etc.)

Title: Collect and Report Information

Level: Basic

Purpose: To give the participant practice in assessing and identifying situations accurately, organizing observations into a report format, and reporting information and communicating on a tactical radio.

Tasks and Activities: Conduct Observation, Identify Enemy Locations, Personnel, Vehicles, Aircraft, Report Information (SALUTE or SALT), Send a Radio Message

Scope: The participant is placed in a tactical or semi-tactical situation (OP, check point, watch tower) with a standard or a simulated radio. He is provided a call sign. He is told to report (all, suspicious, specified) information.

SALUTE is a standardized format for organizing and reporting information (modified to SALT under some tactical situations). Meanings and descriptions:

SIZE: Number of persons, vehicles, aircraft.

ACTIVITY: What the observed was doing.

LOCATION: Grid coordinates or reference (distance and direction) from a known point.

UNIT: Description of clothing, patches, marking, numbering, symbols, distinctive identifiers of the observed.

TIME: Date/time group of observation.

EQUIPMENT: Description or identification of any equipment associated with the activity.

A variety of situations should be available including transporting or varying the presentation of the stimulus. The participant's role is primarily passive; the emphasis is on the observing and reporting rather than the tactical response. Not all situational presentations should be obvious (i.e., an enemy soldier); some should be of "neutral" situations that the participant is expected to assess.

The information is normally transmitted by radio. This should be duplicated, including the response by the receiver (which could include instructions like requests for clarification, more information, continued updates). Call signs should be used.

An example of a complete transmission would be: ROMEO ONE ALPHA THIS IS BRAVO SEVEN TANGO OVER. (Bravo Seven Tango This Is Romeo One Alpha Over) ROMEO ONE ALPHA THIS IS BRAVO SEVEN TANGO MESSAGE FOLLOWS OVER (Bravo Seven Tango This Is Romeo One Alpha Roger Over) THIS IS BRAVO SEVEN TANGO SPOT REPORT. SIERRA THREE PERSONNEL. ALPHA MOVING DISMOUNTED ON TRAIL. LIMA VICTOR KILO ONE ZERO NINER FIVE FIVE ZERO. UNIFORM KRASNOVIAN ARTILLERY FLASH ON HELMETS. TANGO ONE FIVE NOVEMBER ELEVEN HUNDRED ZULU. ECHO ONE RADIO AND THREE RIFLES POSSIBLE LASER DESIGNATOR PACKS OVER. (This is Romeo One Alpha Roger Out)

Simulation Specifics: Needs to present a variety of situations and stimulus (persons, vehicles, aircraft, hidden, partial, etc.) under a variety of conditions (terrain, light). Since much observation is through optics (e.g., binoculars) this would be a nice feature, if doable.

Notes: The radio portion could be monitored or recorded for live interpretation; it need not depend on the simulation since no action is required off of it. This is an individual task rather than a pure 'leader' task but is based on leader requirements to systematically assess and report information. This task becomes increasingly important in OOTW (operations other than war - which is becoming more and more the training focus) where the emphasis is on observing and accurately reporting rather than acting and where not everything observed is an 'enemy' (Examples: Sinai Multinational Force, Macedonia, Haiti, Somalia, Bosnia). The use of simulation has a lot of appeal because it could be specifically tailored to the type of presentations and reporting that fits a specific region, theater, or situation.

Title: Conduct A Dismounted Patrol

Level: Intermediate

Purpose: The focus of the training is on a soldier who is the squad leader of a dismounted standard infantry squad on a patrol mission. A series of events, continuous but separable, are set to occur by controlling the stimulus in the form of the terrain and cover conditions, enemy, and directions given to the squad leader.

Tasks and Activities: Use Visual Signals, Control Movement of a Squad, Move as part of a Squad formation, Conduct Observation, Identify Enemy Locations.

Scope: Initial Instructions: "You are a squad leader conducting a dismounted daylight patrol. Your mission is to reconnoiter this trail from (here) to (here) and make sure it is clear for the rest of the platoon to follow. Enemy activity in the area has been primarily individual snipers and roving groups of two and three man patrols. For the first ___ meters, the area was swept this morning and is clear. Beyond that, you can expect probable enemy contact. You are to practice noise discipline and radio listening silence unless contact is made. You are to complete your reconnaissance NLT ___ ."

Event One: Initial Movement. Special Instructions: "You have already briefed your squad on the initial movement from here to the woodline. You will start off in a diamond formation. You are located at the ___ position. You may change the formation any time you wish. You are to observe noise discipline which means you must use arm and hand signals. The squad will respond as you direct them."

Event One Performance Requirements: Controls formation, direction, distance, speed, and orientation of movement. Signals move out, halt, speed up, slow down, close up, open up as required by the terrain and actions of the squad.

(Movement is over moderately open terrain and there is minimal risk of contact. Intent is to make this event minimally demanding allowing the performer to 'get used' to the situation.)

Event Two: <u>Move by bounds</u>. Special Instructions: " For the next ___ meters, contact is imminent. You should move by bounds. Team Alpha is your initial movement team and Team Bravo is your initial overwatch team. You may position yourself with either team. You must control the activities of both teams"

Event Two Performance Requirements: Moves by bounds, provide overwatch, control movement through arm and hand signals, control length, direction, speed of movement. Control intervals. Position soldiers. Use alternate or successive bounds.

(Movement should require no more than two bounds. No contact. This is more difficult than Event 1 but still easy. There may be a problem with specifying bounding and overwatch control features because this would normally be done with detailed verbal instructions to the two fire teams and may be beyond the simulation capabilities. The event could still be made to happen by specifying that the performer has to go with the movement team and control their movement and automating the overwatch team.)

Event Three: <u>React to Contact.</u> Special Instructions: "The trail enters dense vegetation. Enemy contact is probable, however, you are still under the restrictions of noise discipline and radio silence until contact is made. If contact is made, you may request mortar support through your platoon leader."

Event Three Performance Requirements: Controls formation, direction, distance, speed, interval, positioning, overwatch of movement. Gives arm and hand signals. Reacts to enemy fire by controlling fire and maneuver of squad through arm and hand signals and voice commands. Reports situation. Requests and adjust mortar fire.

(Very close terrain and vegetation limits sight. Enemy opens fire from covered and concealed positions, requiring identification of location, and maneuver supported by fire. Enemy may choose to withdraw after initial attack. This is meant to be the more difficult event but only in terms of the time pressures put on the performer and the conditions of performance. Actual activities required are no harder and differ little from the first two events. Initial part of the event [until contact] is just another movement requirement, only under more difficult conditions. One of the main performance measures is that he assess the situation and accurately report it to the platoon leader; this is not interactive with the simulation but requires him to 'read' the simulation cues. The call for/adjust fire is an option. If it cannot be done, it could be dropped.)

<u>Simulation Specifics:</u> Requires control over types of terrain available as changes in terrain and vegetation dictate changes in formation, intervals, distances, speed.

Requires a semi automated force (SAFOR) for the friendly fire teams that is responsive to the live cues (signals, voice commands, body movements) both individually and as a group, and that can be programmed to execute movements, overwatches, and execute actions on contact. Control over the SAFOR has to be such that it can be programmed to take actions that require corrections by the live performer and the SAFOR must be responsive to those corrections.

OPFOR must be programmable as to activities, lethality, size.

Notes: Would anticipate about a 20 minute requirement for each event. Events should be able to be played back, including the activities of the live performer.

Title: Call For and Adjust Fire

Level: Intermediate

Purpose: To give the participant practice in assessing and calling for mortar or artillery fire in a time pressure, tactical situation.

Tasks and Activities: Identify Enemy Positions and Locations, Communicate Information on Tactical Radio, Call For Fire, Adjust Fire, Read a Map, Determine Azimuth, Sense Indirect Fire.

Scope: Participant is a squad leader in a prepared or hasty defensive position. He will be performing as an indirect fire observer (FO). He has been told he has direct support (either artillery or mortars) and has a radio and has established communications with the FDC. A threat is presented that is appropriate to indirect fire support (either because of size, or could be a vehicle such as a BMP, or could be a situation where smoke for screening is appropriate). If a threat, it could be that they are attacking or fleeing so that if not accurately or timely engaged, bad things happen.

There are some complexities to performing this. Location of the target can be by 3 common methods (grid, polar, shift from known point). Grid is the most common. The observer must determine the location of the target and his direction to the target. He must also give the FDC his location. Also, calls for fire from non indirect fire asset sources often require authentication. (In a training situation, some or all of this information could be 'given' or excluded.)

The initial call for fire is pretty straight forward. It requires the following elements, of which only the first four are standard:

Observer Identification:	Who is requesting fire (call sign)
Warning Order:	How the target is being identified (grid, polar, shift) and what types of fire (adjust, fire for effect, suppress).
Target Location:	Grid, direction.
Description of Target:	Number, type and activity
Method Of Engagement:	Restrictions, type of ammunition (this element is optional)
Method of Control:	At my command , cannot observe (this element is optional)
Authentication:	Standard SOI authentication (this element is optional - but not by the observer)

An example of a complete initial call for fire (and response) would be: MIKE SEVEN FOUR THIS IS QUEBEC ONE ONE ADJUST FIRE OVER (Quebec One One This is Mike Seven Four Adjust Fire Out) GRID XRAY MIKE ONE EIGHT ZERO FIVE ONE THREE OVER (Grid Xray Mike One Eight Zero Five One Three Out) INFANTRY PLATOON IN THE OPEN VICTOR TANGO IN EFFECT OVER (Infantry Platoon In the Open Victor Tango In Effect Authenticate Papa Bravo Over) I AUTHENTICATE CHARLIE OUT

Adjustment of fire is more complex. An initial round is fired. The observer is trying to both bring the round on line (right/left) and on range (over/short) of the target. To do this, he first must sense where the round landed in relation to the target, apply some rules of geometry, and issue a correction. Adjustments are continued with a single round until the round lands within 25 or 50 meters (depending on if it is mortar or artillery) of the desired target point. An initial adjustment would sound like: DIRECTION FIVE TWO ONE ZERO LEFT ONE FIVE ZERO DROP TWO HUNDRED.

Generally an initial call for fire should be made inside of two minutes after target detection, adjustments made (transmission complete) within 30 seconds after round impact, and Fire For Effect issued within nor more than 5 adjustments (somewhat dependent on the nature of the target).

<u>Simulation Specifics:</u> The simulation doesn't have to be capable of 'hearing' the request for fire and adjustment but must be capable of moving fires around as requested. Usually this is done by a human operator. SIMNET has this capability but it doesn't work real well and is difficult for the operators so they use a "bomb button" instead. It is not as accurate. There is no adjustment in SIMNET; if the original fire mission misses, they put in a new mission at a new location. It would be good if the information supplied by the observer (including adjustments) were accurately inputted so that the accuracy and effects could be realistically played. Again, optics are almost always used for this task.

<u>Notes:</u> This is an individual rather then a group task. In theory, any soldier can call for fire. In reality, it is a specialized endeavor; it will be the senior person in the unit in absence of a qualified artillery FO. So the task is very appropriate for E6. Although the <u>procedure</u> is normally trained outside of simulation, the adjustment requirement and the performance under pressure can only be done in some type of simulator. On the other hand, sensing (involving distance estimation and depth perception) is currently well-nigh impossible in a simulation.

<u>Title:</u> Set Up and Occupy Hasty Defensive Positions

<u>Level:</u> Intermediate

<u>Tasks and Activities:</u> Prepare Defensive Position, Establish Perimeter Defense, Set Up OP/LP, Position Weapons, Select Squad Positions, Select Fighting Positions, Select Fields of Fire, Prepare for Attack, Engage Enemy.

Scope: The training participant is a squad leader with a standardized or reinforced infantry squad. He is given an orientation (on a map or on the "ground") to his area and told to set up a hasty defense of a strong point or to establish a perimeter defense. He is given a general area to set up in. He must pick the exact place to defend and position his squad. He must position his SAW in the most likely enemy avenues of approach and position his grenadiers to cover dead space. He must provide for 360 defense and for overlapping fires. He should position Claymores or obstacles in areas he cannot cover. He must position OP and provide for communication or withdrawal. He must provide for alternate and or supplementary individual fighting positions. All positions must provide for cover and concealment. He must plan for and occupy routes of withdrawal from the position and provide for rally points or supplemental squad defensive positions.

The trainee should be given some minimal, but realistic, time to prepare. The time would preclude construction of prepared positions. To 'test' the defense, his position should be attacked by at least a force double that of the squad. Squad members should be capable of becoming causalities.

Simulation Specifics: This may very difficult, if not impossible, to do with SAFOR. There are just too many variables about what individual soldiers have to do on the defense on their own. There are a lot of things the squad leader must direct, and must check, but it is unrealistic that he would perform each. Programming SAFOR to act and react appropriately is probably unrealistic. Therefore the squad may actually have to be manned with real or role-player personnel.

An 'intelligent' automated OPFOR, that could probe and try to find weaknesses, is required.

Should be capable of changing the terrain and the location to provide variety in defensive situations.

Notes: There are same basic, firm principles that squad leaders should be able to apply in setting up a defense. This is a true collective task. There are a lot of sub-activities that need to be applied but the ones that are the squad leader's responsibility are fairly easily identifiable and definable.

Title: Enter and Clear a Building

Level: Advanced

Purpose: To systematically enter, search, and clear a building, destroying all enemy, as part of combat in urban terrain.

Tasks and Activities: Select Covered and Concealed Positions, Provide Overwatch, Enter a Building, Check For Booby Traps, Employ Hand Grenades, Engage Enemy, Use Visual Signals

Scope: The training participants are the squad leader and members of his squad. The building should be at least two stories, with a basement. The squad leader must establish the outside force and the assault force. He must select the entry point, which should be the highest point and avoid obvious entry points like doorways. Ropes, grapples, or rappelling may be required. His force must clear the entry. Inside, they organize into support teams and assault teams. Each hallway and each room must be cleared systematically. Participants must use cooked off grenades and automatic weapons fire in every room. They must employ a variety of methods in entering rooms to avoid a pattern. They must check for, discover, and disarm booby traps. They must clear obstacles. They must keep constant track of each other through voice alerts and announce all entries and exits from rooms and hiding places.

Simulation Specifics: This is DOOM in a training session. It is impossible to do with SAFOR; it requires real people. OPFOR should be automated.

Notes: This is a very difficult task to train in a "real" setting, especially with any kind of opposing force. There is value to this for every member who participates; it is about as close as we come to a truly synchronized performance. All 'leaders' should participate in this task in all positions to truly understand the requirements of this activity which is why it is a good task to include even if it is not trained in a unit setting.

Title: Conduct a Point Ambush

Level: Advanced

Purpose: To select a location to ambush enemy forces, avoid detection, provide early warning, position forces, execute the ambush, destroy all enemy, and escape from the area rapidly and without casualties.

Tasks and Activities: Select an Ambush Site, Prepare an Ambush, Position Forces and Weapons, Avoid Fratricide, Select Covered and Concealed Positions, Avoid Detection, Engage Enemy.

Scope: The training participant is the squad leader. He is oriented to a particular location and told to prepare a point ambush. He is given the expected ambush target (dismounted troops and numbers, vehicles) and may be supplemented with special weapons or munitions (mines, anti-tank, machine-guns). He must select the site for the ambush and identify the kill zone limits. He must establish flank security and provide for early detection. He must set up mines and automatic weapons and grenades to cover the kill zone. He must provide for total concealment. He must position personnel and designate fields of fire to cover the kill zone. He should provide for an assault force. He must institute control measures to control opening, shifting, lifting and cease fires. He must position individuals and institute control measures to avoid fratricide. He

must execute the ambush to maximize the kill. He must withdraw his force rapidly and meet at a preselected point. He must minimize friendly casualties.

A smart OPFOR will try to discover and circumvent or disrupt or even counter-attack. There should be some way of evaluating this in light of the unit's preparation.

Terrain and target changes can provide variety. Ambushes can involve vehicles or urban settings.

Simulation Specifics: Simulation needs to allow for concealment. SAFOR should be possible, but not easy with current state of the art, for the rest of the squad. An automated OPFOR should be possible.

Title: Conduct An Assault

Level: Advanced

Purpose: To practice organization, control, and conduct of dismounted assault on hasty and fortified enemy positions.

Tasks and Activities: Conduct an Assault, Support by Fire, Conduct Fire and Maneuver, Engage Enemy Targets, Use Visual Signals, Give Fire Commands

Scope: Given an identified enemy position appropriate for a squad objective. Training participant is the squad leader. He is located in an attack position short of the objective. He must organize his assault force and his covering force and pick positions for both. He must plan for employment of indirect fires and smoke. He must provide for the lifting and shifting of organic fires and indirect fires. He executes the assault, controlling both the assault forces and the supporting forces. Fratricide is a concern and should be a measurable item.

A small, but well prepared OPFOR is required.

Variety can be provided by terrain and by modifications such as fortified and mined areas, or a bunker, or a building.

Simulation Specifics: SAFOR should be workable for the squad forces. Automated OPFOR should be easy. Cover and concealment in movement routes and cover and concealed individual positions up to the final assault is a requirement. So is the ability to employ smoke effectively. Mines, booby traps, and obstacles should be in place.